TECHNICAL RESEARCH REPORT

Nonlinear Model Reduction for RTCVD

by A. Newman, P.S. Krishnaprasad

CDCSS T.R. 98-10
(ISR T.R. 98-42)

The Center for Dynamics and Control of Smart Structures (CDCSS) is a joint Harvard University, Boston University, University of Maryland center, supported by the Army Research Office under the ODDR&E MURI97 Program Grant No. DAAG55-97-1-0114 (through Harvard University). This document is a technical report in the CDCSS series originating at the University of Maryland.

Web site http://www.isr.umd.edu/CDCSS/cdcss.html

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
1998		**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Nonlinear Model Reduction for RTCVD	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Army Research Office,PO Box 12211,Research Triangle Park,NC,27709**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**7**	

Nonlinear Model Reduction for RTCVD

Andrew J. Newman *
Institute for Systems Research and
Electrical Engineering Department
University of Maryland
College Park, MD 20742
newman@isr.umd.edu

P. S. Krishnaprasad
Institute for Systems Research and
Electrical Engineering Department
University of Maryland
College Park, MD 20742
krishna@isr.umd.edu

Abstract

In this paper, we examine alternative methods for reducing the dimensionality of nonlinear dynamical system models arising in control of rapid thermal chemical vapor deposition (RTCVD) for semiconductor manufacturing. We focus on model reduction for the ordinary differential equation model describing heat transfer to, from, and within a semiconductor wafer in the RTCVD chamber. Two model reduction approaches are studied and compared: the proper orthogonal decomposition and the method of balancing. This leads to a discussion of computational issues in the practical implementation of balancing for nonlinear systems.

1 Introduction

Model reduction deals with methods for reducing the dimensionality of dynamical system models. The motivation is that models of lower dimension are less complex and easier to work with for various purposes such as simulation, optimization, and control. One source of dynamical system models which are good candidates for model reduction is rapid thermal chemical vapor deposition (RTCVD), a process used to deposit thin films on a semiconductor wafer via thermally activated chemical mechanisms.

This paper describes our recent progress and ongoing investigation in nonlinear model reduction for RTCVD. These efforts have taken place as part of a joint industry–academia research project to optimize the epitaxial growth of silicon–germanium (Si–Ge) heterostructures in a commercial RTCVD reactor. Participants are the University of Maryland's Institute for Systems Research (ISR) of College Park, MD, and the Northrop Grumman Corporation's Electronic Sensors and Systems Division (ESSD) of Linthicum, MD. Details and some results from this project are described in [1, 2].

Our approach to model development for control and optimization combines first–principles, experimental and simulation data, and system–theoretic ideas for purposes of model validation and model reduction. This paper focuses on two main topics: a comparison of model reduction approaches as applied to one component of the overall RTCVD model, and computational issues in nonlinear balancing. Attractive properties, drawbacks, and challenges of model reduction via established statistically–based empirical methods and balanced realization methods are discussed, demonstrated, and compared via numerical studies.

*This research was supported by grants from the Northrop Grumman Foundation, the National Science Foundation's Engineering Research Centers Program: NSFD CDR 8803012 and NSF Grant EEC-9527576, and the Army Research Office under the ODDR&E MURI97 Program Grant No. DAAG55-97-1-0114.

2 RTCVD Models

Process and equipment models for RTCVD involve prediction of time evolution for heat transfer within and among the wafer, chamber walls, and process gases (temperature fields in solids and gases), momentum transport in the gas phase (gas flow velocity vectors), mass transport in the gas phase (species concentrations), and chemical reaction kinetics in the gas phase and at the wafer surface (reaction rates, deposition thickness). Thus, they consist mainly of balance equations for conservation of energy (heat equation), momentum (Navier–Stokes), and mass (continuity), along with equations that describe the relevant chemical mechanisms (e.g. Arrhenius laws) and boundary and initial conditions.

In their continuum form, the balance equations for RTCVD yield a system of nonlinear coupled partial differential equations (PDEs) with associated boundary conditions (BCs) and initial conditions (ICs). Lumped versions of the equations can be obtained using an appropriate discretization method, e.g., finite–elements, to yield a system of coupled nonlinear ordinary differential equations (ODEs). The system of ODEs can be decoupled by invoking certain simplifying assumptions. Even the resulting simplified nonlinear system is typically of relatively high–order, so that not only is the model computationally demanding for simulation, but moreover it is computationally prohibitive for real–time control. Thus, the motivation for reducing the model order is apparent.

To illustrate ideas in model reduction, we focus on the evolution system describing heat transfer on the surface of a silicon wafer in the ASM Epsilon–1, a commercial RTCVD reactor used by Northrop Grumman ESSD. The Epsilon–1 has a horizontally oriented process chamber,

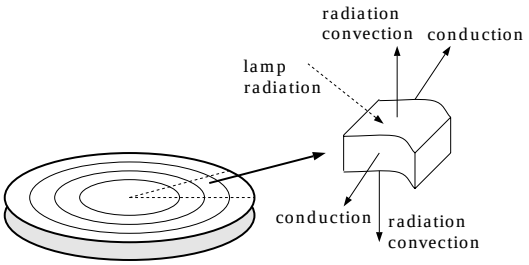

Figure 1: Semiconductor wafer discretized into annular regions. Heat transfer in wafer is described by energy balance for each discrete element.

with gases flowing parallel to and over the top surface of the wafer. The silicon wafer is very thin and rotating on a susceptor so that it is reasonable to assume axial and azimuthal symmetry, i.e., all mechanisms, including heat transfer, are functions only of time t and radial position r on the wafer surface. The wafer is heated by lamps, grouped into "lamp zones", whose radiant intensity profiles are functions of r with corresponding power levels that are available as control inputs. The relevant heat transfer mechanisms are shown in Figure 1.

A simplified version of the temperature field evolution on the surface of a silicon wafer during RTCVD is given by the ODE

$$\dot{T}_w = A_c\, T_w + A_r\, T_w^4 + A_v\, T_w + \Gamma + B\, P \qquad (1)$$

where $T_w(t)$ is a n–vector of temperatures corresponding to radial positions on the wafer surface, A_c, A_r, and A_v are constant $n \times n$ matrices representing the effects of conductive, radiative, and convective heat transfer mechanisms, respectively, Γ is a constant n–vector which accounts for the gas and chamber wall temperature, B is a $n \times m$ matrix of discretized lamp zone radiant intensity profiles, and P is a m–vector of control inputs corresponding to lamp zone power levels. Note that for the nonlinear term T_w^4 the exponent is taken component–wise. Our model uses $m = 3$ lamp zone actuators and $n = 101$ grid points for the spatial discretization. The parameters and lamp zone intensity profiles used in A_c, A_r, A_v, Γ, and B are derived from material properties, radiative heat transfer analysis, and experimental results (see [2]).

To model the measurement of temperature at discrete points on the wafer surface via thermocouples, we augment the nonlinear state equation (1) with the linear output equation

$$T_{tc} = C\, T_w \qquad (2)$$

where T_{tc} is a p–vector of thermocouple measurements, and C is a $p \times n$ matrix with entries corresponding to thermocouple locations. We use $p = 3$ thermocouples in our model, at the center, edge, and midpoint between center and edge.

Later, we make use of a linearized version of (1). The state equation is linearized about the constant temperature profile $T_w = \Gamma$ and is of the form

$$\dot{x} = A\, x + B\, u \qquad (3)$$

with

$$A = A_c + A_v + 4F$$

where

$$[F]_{ij} = [A_r]_{ij}\, \Gamma_j^3$$

and x and u are translations of T_w and P, respectively.

3 Model Reduction

A general approach to model reduction is to find a coordinate transformation of the state space under which the state components can be ranked in a meaningful way in terms of their influence on the system behavior. Then, state components of the transformed system with relatively small influence can be truncated without substantially degrading the correctness, i.e., predictive capability, of the model. We note that for systems evolving on $\mathbb{R}^n$, each coordinate transformation can be identified with a corresponding set of basis n–vectors.

Proper Orthogonal Decomposition

One approach to finding a basis for the desired coordinate transformation is application of the Karhunen–Loeve decomposition (see, e.g., [3, 4, 5, 6]), also known as the proper orthogonal decomposition (POD), method of empirical eigenfunctions, or principal components analysis (PCA). The POD is a statistical pattern analysis technique for finding the dominant structures in an ensemble of spatially distributed data. These structures can be used as an orthogonal basis for efficient representation of the ensemble. In particular, the POD basis elements are the orthogonal eigenfunctions of the two–point spatial covariance of the data ensemble. When the data ensemble consists of vectors in $\mathbb{R}^n$, the POD basis vectors are just the the columns of the matrix U in the singular value decomposition

$$X = U \Sigma V^T \qquad (4)$$

where X is a matrix whose columns are the members of the data ensemble.

For the case of a dynamical system describing a flow, e.g., a heat flow, the data ensemble is typically time series data, i.e., "snapshots" of the flow captured at equally spaced intervals in time. The POD coordinate transformation diagonalizes, or decouples, the covariance of the time series data. The basis elements are the principal axes of the flow which generated the time series empirical data. Each has a corresponding eigenvalue (given by the entries on the diagonal of Σ in (4)), which provides a measure of the relative "energy", i.e., mean square fluctuation, associated with that particular direction in the state space. This measure can also be interpreted as the relative amount of time that the flow spends along the corresponding principal axis. Thus, it serves as a well–defined measure of the influence of a state component on the system behavior.

The POD basis is optimal from the points of view of data compression and error minimization. Specifically, a

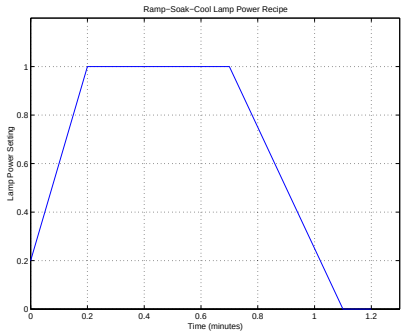

Figure 2: Lamp power settings for RSC recipe.

truncated series representation of the data has a smaller mean square error than a representation by any other basis of the same dimension (see, e.g., [3]).

Application

The attractive properties of the POD have led to success in applying it to areas such as turbulence modeling (e.g., [3]) and pattern recognition (e.g., [7]). Recently, much work has been done to study its use for RTCVD model reduction (e.g., [8, 9, 10, 11]). We have performed a similar study, applying the POD technique toward finding a low–order approximation to (1), described as follows.

To generate empirical time series data, i.e., snapshots of the wafer temperature field, the system (1–2) was simulated using two different types of control input recipes. They are referred to as Ramp–Soak–Cool (RSC) and Perturbation–Of–Constant (POC).

The RSC recipe mimics a typical processing recipe in which a lamp zone power setting is gradually ramped up from zero to full power, maintained at full power for a specified period of time, and then gradually ramped down from full to zero power, as shown in Figure 2. This recipe is applied to one of the lamp zones, while the other two zones are held at zero power. The simulation is then repeated using RSC individually for each of the other two lamp zones. The entire ensemble (three sets) of time series data is combined and used to compute the POD basis elements, which are then ranked according to magnitude of associated eigenvalue. The basis elements with the four largest eigenvalues are shown in Figure 3. Corresponding relative energy values are contained in Table 1.

The POC recipe applies small perturbations of a predetermined set of constant power settings which, if left unperturbed, would result in a uniform steady state temperature field of 1000K. The perturbations are achieved by adjusting the power setting of each lamp zone, one at a time, first to 110% and then to 90%, of the original setting. This results in 6 different control recipes (see [2] for details).

The system response to excitation from each of the six POC recipes is sampled and combined as the time series data for computing POD basis elements. Once again,

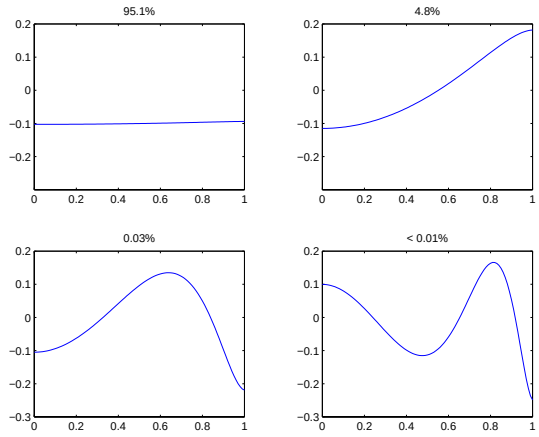

Figure 3: Basis elements computed using POD from RSC empirical data.

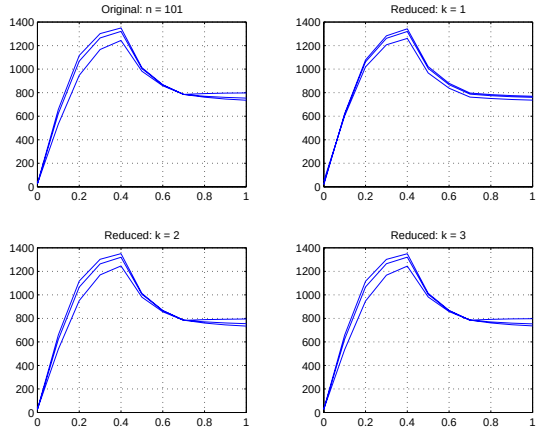

Figure 4: Thermocouple readings for original and reduced models with Test Recipe 2 using POD RSC coordinate transformation. Original $n = 100$. Reduced $k = 1, 2, 3$.

basis elements are ranked by corresponding relative energy value, contained in Table 1.

We now test the efficiency of each POD basis by projecting (1) onto a k–dimensional subspace spanned by the k most important basis vectors using standard Galerkin projection, where $k \ll n$. The reduced order approximation is numerically integrated using two test recipes as control inputs, that are different from the recipes used to generate the RSC and POC data ensembles.

Test Recipe 1	$P = [0.5\ 0.5\ 0.5]$	$0 < t \leq 60$
Test Recipe 2	$P = [1.0\ 0.0\ 0.0]$	$0 < t \leq 24$
	$P = [0.0\ 1.0\ 0.0]$	$24 < t \leq 42$
	$P = [0.0\ 0.0\ 1.0]$	$42 < t \leq 60$

Figure 4 shows simulated thermocouple readings for Test Recipe 2, where the k basis vectors for the Galerkin projection are from POD RSC, for $k = 1, 2, 3$. For complete results see [2]. The error between original and reduced models is computed as the maximum deviation between actual and approximate thermocouple readings. Results for all test simulations are contained in Table 2.

Shortcomings

It is clear that the efficiency of a basis determined via the POD method depends strongly on how well the data ensemble captures the relevant system behavior. This leads to serious shortcomings for model reduction of control systems with inputs and outputs. The POD basis elements will be sensitive to the choice of control inputs, ICs, and BCs used to generate the empirical time–series data. For nonlinear systems, small perturbations in these parameters may produce qualitatively different system responses. In addition, the data may fail to capture dynamical effects occuring at widely differing time scales. Usually, these issues are ignored, because the model is only being used for a particular purpose, e.g., to simulate tracking of one particular reference trajectory. In that case, a representative set of control inputs, BCs, and ICs is selected for generating the data ensemble. But the optimality property of the POD basis, and the predicitive capability of the resulting reduced order model, may be localized to a relatively small region of the space of allowable inputs, BCs, and ICs. In addition, the POD approach fails to consider the influence of state components on the system measurements, or outputs. It would be desirable to truncate state components that have small influence on the outputs, i.e. that do not appear in our measurements. The POD method does not do this, thus its efficiency may be diminished when constructing "black–box" input–output models of the system.

Balancing

In response to these and other concerns, we considered an alternative approach based on the method of balanced realizations (see, e.g., [12, 13]). In this method, a coordinate transformation is computed which allows state components to be ranked according to their influence on the input–output behavior of the system as measured by the Hankel norm of the system, i.e., the gain from past inputs to future outputs. The resulting basis for the linear transformation makes the transformed realization "equally controllable and observable," i.e. balanced, and depends only upon intrinsic properties of the original model, specifically controllability and observability, embodied in its evolution equation and output equation. In the linear case, explicit bounds can be computed for the error between the original and reduced order models. These error bounds are independent of any particular set of control input signals, BCs, or ICs. Although explicit error bounds have not yet been found for the nonlinear case, we still wish to exploit the property that the correctness of an approximation using a truncated balanced realization does not depend upon generating a representative data ensemble. Furthermore, the balancing method emphasizes state components that are both strongly controllable and strongly observable, so that state components which are least likely to influence the measurements are truncated.

Application

We have applied the balancing method to the linearized control system given by (3) and (2). The linearized version is used because efforts toward development of useful algorithms for nonlinear balancing are still underway, as described later in this paper.

Numerical difficulties arise when applying standard linear balancing algorithms to our linearized system. In particular, the system is nearly nonminimal, i.e., the condition numbers of the controllability and observability Gramian matrices are very large (see [2] for details). Safanov and Chiang [14] have proposed a method for overcoming such difficulties, based on the ordered Schur decomposition of the product of the controllability and observability Gramian matrices. We have used their method to produce coordinate transformation basis vectors and kth–order reduced models for the linearized system. The results are shown in Tables 1 and 2.

Discussion

Due to the shape of the lamp zone heat flux intensity profiles and the smoothing effect of the diffusion operator, the evolution of the wafer temperature field does not produce especially interesting behavior, e.g., spatial profiles whose fluctuations from the mean vary substantially in the mean square sense from the initial profile, assuming the initial profile is relatively smooth. Thus, we expect little difficulty in capturing the essence of the input–output behavior of the system in a low dimensional model. Our results show that this is indeed the case.

For the inputs used in the validation tests, the input–output behavior of the wafer heat transfer system can be reconstructed using reduced models of order 4 so that thermocouple readings are within 1 degree C of the readings obtained using the original model. This holds whether the POD or balancing method is used, and for whichever empirical data ensemble was used for computing the POD transformation. Even reduced models of order 2 produce a reasonable approximation with "worst case" errors less than 15 degrees C.

The POD method appears to have performed slightly better than the balancing method in this study. One reason for this result is that the balancing transformation was computed for the linearized system, while the validation tests were performed for the reduced order nonlinear system. Another reason is the simple input–output behavior of this particular system. The principal components of the flow are relatively insensitive to the choice of inputs, and hence, any set of empirically determined eigenfunctions for this system will likely be efficient for model reduction purposes.

Thus, the results of this preliminary study are not decisive regarding choice of reduction method. We seek a model reduction approach which provides explicit error bounds, does not depend on a specific data ensemble, and which can be applied directly to the nonlinear control system.

Method	Mode 1	Mode 2	Mode 3	Mode 4
POD RSC	95.06	4.77	0.14	0.03
POD POC	93.43	6.25	0.23	0.09
Balancing	98.02	1.83	0.13	0.02

Table 1: Normalized eigenvalues, i.e., percent energy, corresponding to basis elements used in model reduction for POD method with RSC data, POD method with POC data, and balancing approach.

Simulation	Reduction Method	Reduced Model Order 1	2	3	4
Test 1	POD RSC	27.23	2.68	0.58	0.11
	POD POC	26.85	1.26	1.13	0.10
	Balancing	50.68	7.03	0.44	0.08
Test 2	POD RSC	72.33	5.22	1.48	0.18
	POD POC	72.60	4.79	4.35	0.43
	Balancing	80.81	14.28	1.70	0.12

Table 2: Maximum deviation (degrees C) between outputs of original and reduced models for POD method with RSC data, POD method with POC data, and balancing approach.

4 Computational Issues In Nonlinear Balancing

In order to address some of the deficiencies in model reduction via the POD and linear balancing approaches, we seek a procedure to apply the balancing method directly to the nonlinear system of interest. We have already seen that for linear systems, the balancing coordinate transformation can be determined efficiently using known algorithms. The theory of balancing was extended to a class of nonlinear systems by Scherpen [15], in which a general theory and procedure is proposed. In contrast to the linear case, the theory and procedure for nonlinear balancing is not immediately amenable to computational implementation. Here, we begin to address some important computational issues in balancing for nonlinear systems.

We consider the balancing procedure presented in [15] for open loop stable nonlinear systems of the form

$$\dot{x} = f(x) + g(x)\,u \tag{5}$$
$$y = h(x) \tag{6}$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ are local coordinates for a smooth state manifold denoted by M, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$, f and h are of class C^∞, $f(0) = 0$, and $h(0) = 0$.

To determine the balancing state space transformation we seek a change of coordinates under which the transformed system realization is equally controllable and observable. The degree to which a system realization is controllable or observable, respectively, can be measured in a precise way using the controllability and observability energy functions of the system [15].

Definition 4.1 *The controllability function, $L_c : \mathbb{R}^n \to \mathbb{R}$, and observability function, $L_o : \mathbb{R}^n \to \mathbb{R}$, for (5–6) are defined by*

$$L_c(x_0) = \min_{\substack{u \in \mathcal{L}_2(-\infty, 0) \\ x(-\infty) = 0 \\ x(0) = x_0}} \frac{1}{2} \int_{-\infty}^{0} \|u(t)\|^2 \, dt \tag{7}$$

and

$$L_o(x_0) = \frac{1}{2} \int_0^\infty \|y(t)\|^2 \, dt\,, \tag{8}$$
$$x(0) = x_0\,, \quad u(t) \equiv 0\,, \quad 0 \le t < \infty.$$

In the linear case, the controllability and observability functions specialize to the quadratic forms

$$L_c(x_0) = \frac{1}{2}\, x_0^T\, {W_c}^{-1}\, x_0 \tag{9}$$
$$L_o(x_0) = \frac{1}{2}\, x_0^T\, W_o\, x_0 \tag{10}$$

where the constant $n \times n$ matrices W_c and W_o are the familiar controllability and observability Gramian matrices, respectively.

It is not surprising, then, that a key step in the nonlinear balancing procedure of [15] is to find a change of coordinates under which L_c is locally quadratic in a neighborhood $\mathcal{U}$ of its critical point 0. This is accomplished by appealing to the Morse Lemma (see, e.g., [16]).

Theorem 4.2 (Morse Lemma) *Let p be a non–degenerate critical point for the real–valued function F with index r. Then in a neighborhood $\mathcal{U}$ of p there exists a local change of coordinates $z \mapsto x = \Psi(z)$, $\Psi(0) = p$, such that*

$$F(x) = F(\Psi(z)) = F(p) - \sum_{i=1}^{r} z_i^2 + \sum_{i=r+1}^{n} z_i^2 \tag{11}$$

holds for all $x \in \mathcal{U}$.

There are various difficulties in the computation of L_c and L_o, and numerical implementation of the Morse Lemma. We are currently investigating several approaches for each problem. Here, we discuss the problem of diagonalizing a particular x–dependent symmetric matrix on a neighborhood of 0, which arises in determination of a smooth Morse coordinate transformation for a function F around a critical point 0. The key is that the diagonalization must be done in a smooth way on $\mathbb{R}^n$, and thus results in a search over smooth functions taking values in the matrix Lie group $SO(n)$.

It is known (see, e.g., [16]) that for F with critical point 0 there exists a symmetric matrix $H(x)$ for each x in a neighborhood $\mathcal{U}$ of 0 such that

$$F(x) = x^T\, H(x)\, x \tag{12}$$

with $H(0) = \partial_x^2 F(0)$. Suppose we have an algorithm that produces the desired $H(x)$ for each x. The next step in implementation of the Morse Lemma is to find a smooth diagonalization of the matrix $H(x)$, i.e., find a smooth function $U(x)$ taking values in $SO(n)$ such that

$$U^T(x)\, H(x)\, U(x) = \Lambda(x) = \mathrm{diag}\{\lambda_1(x), \ldots, \lambda_n(x)\} \tag{13}$$

for each x in $\mathcal{U}$.

We propose the following approach. Choose a basis $\{\alpha_1, \alpha_2, \ldots\}$ for the $\mathcal{L}_2$–completion of $C_0^\infty(\bar{\mathcal{U}}, \mathbb{R})$, i.e., smooth real-valued functions with compact support on the closure of $\mathcal{U}$, and a basis $\{A_1, \ldots, A_k\}$ for $so(n)$, i.e., skew matrices, with $k = n(n-1)/2$. Then a Nk–dimensional approximation to the desired $U(x)$ is given by

$$\widehat{U}(x) = \exp\left(\sum_{i=1}^{k} A_i \sum_{j=1}^{N} c_{ij}\, \alpha_j(x)\right) \tag{14}$$

and can be found by solving the finite–dimensional minimization problem

$$\min_{\substack{c_{ij} \\ i=1,\ldots,k \\ j=1,\ldots,N}} \int_{\mathcal{U}} \|\widehat{U}(x)\, H(x)\, \widehat{U}(x) - \Lambda(x)\|\, dx. \tag{15}$$

5 Conclusion

We have presented a comparison of nonlinear model reduction methods, POD and balancing, which appear to perform well for a specialized problem in RTCVD, but give no guarantees of satisfactory performance in general for dimensionality reduction of nonlinear systems with control inputs and outputs. We address these deficiencies by appealing to the theory of nonlinear balancing, and point out difficulties in terms of numerical implementation. We focus on one step involved in computation of the Morse coordinate transformation, in which a smooth $SO(n)$–valued function must be determined to diagonalize a particular x–dependent symmetric matrix. We suggest an approach which requires solution of a finite–dimensional minimization problem.

References

[1] A. Newman, P. Krishnaprasad, S. Ponczak, and P. Brabant, "Modeling and model reduction for epitaxial growth," in *Proceedings of the SEMATECH AEC/APC Workshop IX*, SEMATECH, Sept. 1997.

[2] A. Newman, P. Krishnaprasad, S. Ponczak, and P. Brabant, "Modeling and model reduction for control and optimization of epitaxial growth in a commercial RTCVD reactor." To be published as Technical Report, Institute for Systems Research, 1998.

[3] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems, and Symmetry*. Cambridge University Press, 1996.

[4] M. D. Graham and I. G. Kevrekidis, "Alternative approaches to the Karhunen-Loeve decomposition for model reduction and data analysis," *Comp. Chem. Engr.*, vol. 20, no. 5, pp. 495–506, 1996.

[5] L. Sirovich, "Turbulence and the dynamics of coherent structures; Part I: Coherent structures," *Quarterly Appl. Math.*, vol. 45, pp. 561–571, Oct. 1987.

[6] A. J. Newman, "Model reduction via the Karhunen-Loeve expansion Part I: An exposition," Tech. Rep. T.R. 96-32, Inst. Systems Research, Apr. 1996.

[7] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *Journal of the Optical Society of America A*, vol. 4, pp. 519–524, Mar. 1987.

[8] H. Aling, S. Banerjee, A. K. Bangia, V. Cole, J. Ebert, A. Emami-Naeini, K. F. Jensen, I. G. Kevrekidis, and S. Shvartsman, "Nonlinear model reduction for simulation and control of rapid thermal processing," in *Proceedings of the American Control Conference*, pp. 2233–2238, 1997.

[9] H. Aling, J. Ebert, A. Emami-Naeini, and R. Kosut, "Application of a nonlinear model reduction method to rapid thermal processing reactors," in *Proceedings of the IFAC World Congress*, 1996.

[10] A. J. Newman, "Model reduction via the Karhunen-Loeve expansion Part II: Some elementary examples," Tech. Rep. T.R. 96-33, Institute for Systems Research, Apr. 1996.

[11] A. Theodoropoulou, R. A. Adomaitis, and E. Zafiriou, "Model reduction for optimization of RTCVD systems," Tech. Rep. T.R. 96-64r1, Institute for Systems Research, Nov. 1996.

[12] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞–error bounds," *International Journal of Control*, vol. 39, no. 6, pp. 1115–1193, 1984.

[13] B. C. Moore, "Principal component analysis in linear systems: Controllability, observability, and model reduction," *IEEE Transactions on Automatic Control*, vol. 26, pp. 17–32, Feb. 1981.

[14] M. G. Safonov and R. Y. Chiang, "A Schur method for balanced truncation model reduction," *IEEE Trans. Automatic Control*, vol. 34, pp. 729–733, July 1989.

[15] J. Scherpen, "Balancing for nonlinear systems," *Sys. Cont. Lett.*, vol. 21, pp. 143–153, 1993.

[16] J. Milnor, *Morse Theory*. Princeton University Press, 1969.